what the informants are discussing and how they make sense of it all. And this in turnis embedded in a cultural context, inextricably shaped by forces such as the mass media.

Introspection and careful analysis, while arguably necessary for all disciplines, is evenmore essential for public historians because they are part of the public discourse and their work can have significant social and political repercussions. The interpretation ofevents is not simply a matter of who has the more accurate version. It can also demonstrate who exercises political power and how that power can be abused. A goodexample is the story of Luigi Trastulli, a 21-year old steelworker from Terni (an industrial town in central Italy), who died on March 17, 1949 after clashing with the local police. The actual incident was minor, compared to the high-casualty police violence of the 1950s, but the way the event was appropriated was significant. It "became the ground upon which collective memory and imagination built a cluster of tales, symbols, legends, and imaginary reconstructions."

The different ways in which this particular event was recalled tells volumes about the power relationship between the workers and the police. The official account made by the police stated that the steelworkers went on a major strike without having obtained permission for it, and that the violence started with the mob. Workers who were interviewed recalled on the other hand that it was no strike at all, instead several workers were just getting off work leaving the factory at the same time, and that the police initiated the violence. In this instance recall and political power are inextricablybound and cannot be separated. The actors' state of mind at that time helped determine how the event was interpreted and what it meant.

Portelli who interviewed workers several decades after the event found that Trastulli'sdeath was later even connected to events that had nothing to do with the police clash in which he died. The protest in 1949 was directed against NATO but many workers came to associate his death with the massive labor strikes of the 1950s. Consequently Trastulli was remembered as a martyr who died during worker strikes in the cause of labor. When Portelli brought this discrepancy to the attention of his informants, they seemed little perturbed. After all to them it made more sense to think of Trastulli's death in the context of labor strikes since these had a more profound effect on the community than the NATO protest several years earlier. Chronology thus seemed lessimportant than the meaning which the workers bestowed and in this respect memory as history can be seen in terms of the symbolic and psychological.

The discrepancy between the official accounts and the recollections of the workers points to another dilemma: the tension between public life and private memory. Classappears to be particularly pertinent here because people's accounts are evaluated according to their social position. While lower-class people may be invited to share

their experiences, they are seldom asked for an analysis of the events. Such discrimination is standard practice and "reinforces the already deeply-rooted, class- based ideology that sees ordinary people as sources of data, rather than as shapers and interpreters of their own experience."

Class is particularly important if we try to get a view from the bottom up. Some people speak anonymously and their experiences are merely historical in the sense that they reflect many experiences. This seemed to be the case with numerous people Studs Terkel interviewed on their experiences during the Great Depression. Many of his informants interpreted the depression in terms of personal failure. These testimonies of survivors loaded with the scars of psychic suffering give great insight into why there has not been a more critical assessment of the Great Depression, why itdid not produce "more focused critiques of American capitalism and culture, more sustained efforts to see fundamental change." As time passes, these experiences of personal failure are seen through the lenses of success and survival, a legacy for the generations that were to follow. A structural criticism was bypassed and history was thus filtered twice, "by time and subsequent experience."

Not all of the survivors of the Great Depression spoke anonymously, though. Some had a direct impact on historical developments and their involvement was clearly in the public realm. Most of the time, the "spectrum of private and public experiences and subjectivity parallels the spectrum of power and position," and those who exertedlittle political power were most likely to interpret the depression primarily as a privatetribulation. Nonetheless there were exceptions, and many of the rich could live quite privately, while some among the poor were politically active and functioned in "a public dimension well beyond their own subjectivity." The kind of information obtained by interviews can thus range from the particular to the general, from specificprivate experiences to a discussion of what it meant—it can be highly anecdotal or fairly analytical.

For the oral historian evaluating what people are saying is perhaps the most difficultaspect of the inquiry but it is also the most important. Here it becomes pertinent to understand how people make sense of their experiences and how they construct meaning out of them. By looking at this "oral history reveals patterns and choices, that, taken together, begin to define the reinforcing and screening apparatus of the general culture, and the ways in which it encourages us to digest experience." Attitudes and opinions that are formed in a cultural context will emerge from these interviews.

However, oral history is not just employed to detect the formation of public opinion, itis also used to shape them. Frisch points out that this was the case with the PBS seriesVietnam: A Television History where primarily those individuals who held powerful

positions were prompted to interpret historical events. Peasants and foot soldiers were interviewed for their personal experiences but not asked to put those into a larger context. “The higher or more important the position of the subject, however, the morelikely he or she is to be seen offering historical judgments of a broader nature, sweeping evaluations of what an event meant, what caused it, or what the public felt.”To Frisch, even more disturbing than this selectivity regarding interpretation is the fact that personal experiences are often presented without ever being put into their historical context. Vietnam is not uncommon; more recent documentaries also fail when it comes to “preventing a momentarily glimpsed reality from slipping back behind a curtain of amnesia, or from receding into a blurry distance.” The problem is that by presenting events such as the Vietnam War without ever seriously discussing the underlying causes, society can maintain its denial and disengagement.
Consequently it can avoid taking responsibility.

In projects such as historical documentaries, oral history can be used for political purposes and cultural hegemony. “In the political arena, for instance, where major conflicts in a democratic society are presumably engaged, the war and its roots were never legitimately discussible.” The causes and ideological underpinnings could not be debated while the war was still going on because this was considered to underminethe war effort. Once the war was over the nation supposedly embarked on a healing process and in this postwar environment a discussion of the war was also treated as subversive. Consequently dissenting voices were silenced, or at least given the least amount of exposure possible. The prevalent manner of dealing with these issues was not to engage in a multi-dimensional discussion but instead to package this complex historical experience and market it through the mass-mediated popular culture for “acceptable public remembering.”

The issue of public memory has also been contentious in Germany, where people are still coming to terms with the Nazi era and the Holocaust. Even today, more than half a century after World War II, the question of responsibility is still controversial and hotly debated by academics. Initially at war’s end the German nation was preoccupiedwith reconstruction and there was little public debate on the Nazi legacy. A massive discussion in the public sphere was not initiated until the television series Holocaust premiered in 1979. Public reaction was strong and interest in the Holocaust peaked, only to ebb shortly afterwards. This short-lived engagement with the issues implies that “the intensity of measurable reaction is a highly questionable indicator of the quality of ‘mass’ memory work.”

It may be difficult to gauge the final impact a historical documentary will have but inGermany oral history has made an important contribution to the debate. One of the driving forces, through which much oral history has been introduced, is the history of everyday life (Alltagsgeschichte), which has been the most important

historiographical development in Germany during the last two decades. Its emphasis has been on everyday, ordinary people, in German called kleine Leute, which can be translated to mean little people, or small people—a term "as suggestive as it is imprecise." Historians of everyday life (Alltagshistoriker) intend to dispel the misconception that the common people have little to contribute to the nation's historical narrative. As their Anglo-American counterparts they stress subjectivity andexperience, as well as the social construction of meaning.

Alltagsgeschichte has many of the same goals Frisch identified for oral historians: a willingness to examine history from the bottom up, to give voice to those who have traditionally been overlooked and consequently to democratize the writing of history. In Germany, however, advocates of Alltagsgeschichte operate in a political and academic environment that has been resistant to innovative approaches. The struggleshave been hard fought and after two decades Alltagshistoriker have made small "inroads into the institutional centers of the profession" even though they enjoy considerable recognition abroad. A discussion about the pertinence of oral history andeveryday life has been highly politicized in Germany. Consequently when Alfred Lüdtke, one of the leading advocates of Alltagsgeschichte, speaks about its import, heis arguing from the position of a dissenter.

Writing history from the bottom up is particularly contentious because it supposedly threatens the academic establishment, in particularly social-science history.
Alltagsgeschichte emerged in the mid-1970s as a dissenting movement from the left and gained momentum during the 1980s. Actually it had much in common with social-science history and the so-called Bielefeld school. Both were committed to coming to terms with the Nazi past and both criticized the "traditionalist methodologyand intellectual outlook of the established West German profession," meaning they opposed the older historicist methods that focused on politics and diplomacy. The Bielefeld school, however, sought to explain history by analyzing society and history through "big structures, large processes, huge comparisons" and this is where Alltagsgeschichte diverges. Its goal was to develop a qualitative understanding by
examining the material conditions of everyday life and by "entering the inner world ofpopular experience" in those contexts which have been considered the cultural domain. For the social-science historians embracing cultural history was clearly an abandonment of scientific principles and over this issue the two camps have exchanged words. Lüdtke and his colleagues have consequently been on the defensiveand have argued that they do not intend to supplant the structural study of social change but try to transcend the "sharp dichotomy opposing objective, material, structural, or institutional factors to subjective, cultural, symbolic, or emotional ones."

For Alltagshistoriker studying the lives of ordinary people does not mean that politicsis left out; in fact their lives cannot be separated from the larger questions of power

and appropriation. History from below sheds light on the lives of “historical losers” and perhaps more importantly demonstrates the internal costs of social transformation.Again, it is a question of political power and how this comes to bear on the lives of everybody. Microhistory has the potential to decentralize analysis and interpretation and it also adds qualitatively to our historical understanding because it captures the ambiguities and contradictions of behavior in ways structural analysis is not able to do.

In Germany the barefoot approach has taken history beyond the institutional confines of scholars and scientists. The public has been involved in discussions regarding museum exhibits, memorial sites and the National Memorial Day (Volkstrauertag), which have been contentious issues because they involve national remembrance of theHolocaust and the Third Reich. The public’s involvement in these negotiations is “intended to make it more difficult to repress and forget, to sweep the past under the public carpet.”

When it comes to debates of such magnitude, it seems necessary to involve the public.This, however, can also be potentially hazardous. In the case of Germany, should modern-day Nazis be allowed to participate in the discourse and to what degree should their views be given credence? It can be argued that their opinions need to be made public to alert people to the dangers of such rhetoric. On the other hand, increased exposure may have the opposite effect and attract some people to their cause. Silencing the Nazis is certainly within Germany’s constitutional rights (even arguably a constitutional obligation) but critics charge that such efforts are repressive.

Perhaps a less stark example is that of complicity during the war. How guilty were those millions of Germans who did not participate in the resistance? And can millionsof German civilians be portrayed as victims? The memories of the survivors have puta human face on these issues and once their personal suffering is acknowledged, it becomes increasingly difficult to speak of national responsibility.

Theoretically it is essential to include the voices from below, at least if we seek to approach the most accurate reconstruction of historical events possible. Yet, what exactly does it mean to incorporate the views from the common people? How is this authority transferred and what are historians giving up? In this respect there are no clear answers and these problems need to be solved as specific projects arise. A German historian who deals with issues of genocide will likely arrive at a different answer than a labor historian who examines workers’ lives under particular conditions. As such there are general theoretical guidelines for oral historians to consider, and thinking about the potential societal impact of one’s work is, in my view, a necessity. The detail, however, has to be worked out in the specific context inwhich the work is done. Method is, undoubtedly, an important consideration but not

the preliminary one. More importantly, in the words of Ronald Grele, is "the mind revealed through conversation." And in this respect the oral historian is as much partof the unfolding story as the informants whose experiences he or she seeks to incorporate into the historical narrative.

Last Updated: 8/12/16

3: How to Read 22,198 Journal Articles: Studying the History of German Studies with Topic Models

Allen Beye Riddell

IN THE PAST DECADE, research libraries have digitized their holdings, mak- ing a vast collection of scanned books, newspapers, and other texts conveniently accessible. While these collections present obvious oppor- tunities for historical research, the task of exploring the contents of thou- sands of texts presents a challenge. This chapter introduces a family of methods, often called topic models, that can be used to explore very large collections of texts. Researchers using these methods may be found not only in computer science, statistics, and computational linguistics but also increasingly in the human and social sciences in fields such as women's history, political science, history of science, and classical studies.[1] This introduction uses a topic model to explore a particular corpus, a collec- tion of 22,198 journal articles and book reviews from four US-based German studies journals: *The German Quarterly*, *New German Critique*, *German Studies Review*, and *Monatshefte*. As this is the first time this cor- pus has been explored using quantitative methods, this introduction also presents a new perspective on the disciplinary history of German studies.

This chapter has three parts. First, I review existing methods that researchers, often historians, have used to explore very large collections of texts. Then I introduce a topic model—a probabilistic model of words appearing in a collection of texts—as an alternative way of reading a cor- pus. I aim to show that a topic model of the German studies journals reveals disciplinary trends that would be immensely time consuming to document otherwise. Finally, I discuss prospects for using topic mod- els in nineteenth-century research generally and in intellectual history specifically.

Existing Approaches: Direct and Collaborative Reading

The early 2000s witnessed the emergence of several library digitization efforts (Open Content Alliance and Google Books, to name two exam- ples). During this period, observers asked what historians might plausibly

do with such vast digital collections. Gregory Crane, a classicist and editor-in-chief of the successful Perseus Digital Library, put the question succinctly in 2006, asking, "What do you do with a million books?"[2] As a practical matter, however, Crane might as well have asked what to do with a thousand books, since carefully reading a thousand volumes already involves more time than many researchers are willing to devote to a single project.

For the sake of brevity, I will refer to any collection of texts as a *very large collection* if it contains more texts than a single researcher would be expected to digest in a year's worth of dedicated reading; 22,198 journal articles would count as a very large collection, as would the proceedings of the British Parliament in the nineteenth century, or all articles pub- lished in an established regional newspaper.[3] What options are available to researchers interested in such collections? If they look to past efforts, they have two strategies available: *direct reading* and *collaborative reading*.

Direct reading is familiar. Regardless of the size of the corpus, researchers may invest the required time to read and digest its contents. There are many examples of scholars reading through enormous collec- tions of texts in the course of their research. The American historian Laurel Thatcher Ulrich spent years reading and rereading the nearly 10,000 diary entries of Martha Ballard, a midwife in Maine around 1800.[4] Examples of studies requiring extensive reading from German cultural and intellec- tual history include Fritz Ringer's *The Decline of the German Manda- rins*, which involved his reading a significant fraction of all books written between 1890 and 1933 by German full professors in the human sci- ences, and Kirsten Belgum's *Popularizing the Nation*, which took among its objects ca. 2,500 issues of the weekly magazine *Die Gartenlaube* (The Garden Bower) printed between 1853 and 1900.[5] Familiarity with a very large collection may also be gained over the course of years of research and teaching. There are many scholars of the nineteenth-century Euro- pean novel—such as Katie Trumpener or John Sutherland—who, I sus- pect, have read a significant fraction of all European novels published in the eighteenth and nineteenth centuries.

A second option, collaborative reading, involves dividing up the task of reading among a number of participants. This approach brings with it the challenge of coordinating among readers. There are many examples of this approach.[6] One effort that managed the problem of coordina- tion particularly well is the Genre Evolution Project, led by Carl Simon and Eric Rabkin at the University of Michigan.[7] Simon and Rabkin gath- ered a team of faculty, graduate students, and undergraduates together to read the ca. 2,000 short stories published in major US science fiction magazines between 1929 and 1999. The team was interested in studying how the science fiction genre changed over time and in testing existing claims about the genre against the evidence provided by the short stories

corpus. No participant read all the stories, but participants did overlap in their reading assignments. To coordinate their efforts, the team focused on gathering information about a range of discrete "features," including the genders and ages of authors as well as characteristics of the narratives, such as whether a story was set in the past or whether uses of technol- ogy led to a "bad outcome." As each story was read by at least two par- ticipants, any reader's judgment could be checked against the readings of others. In this fashion, cases of disagreement could be identified and discussed. In the social sciences, this kind of checking is known as assess- ing interrater reliability.

Another example of collaborative reading is Larry Isaac's study of the "labor problem novel" in nineteenth- and early twentieth-century Ameri- can fiction.[8] Isaac considers a novel a labor problem novel if it contains one of four specific representations of labor union activity (typically, a labor strike). The time frame for his study covers nearly fifty years, from 1870 to 1918. Since thousands of novels were published in the United States during this period, reading through all them for mention of a strike would have been an epic undertaking. Instead, Isaac made use of existing studies and bibliographies of novels from the period and divided up the task of reading candidate labor problem novels between himself and grad- uate students. His team eventually arrived at a list of around five hundred novels fitting the definition.

Both direct reading and collaborative reading may be combined with random sampling. If researchers are interested in investigating trends in book publishing in France between 1800 and 1900 and they happen to have a list of publications from the period, then they may take a random sample and work with that corpus. If the sample is random and suffi- ciently large, the researchers may be confident that significant trends in the larger body of books will be identifiable in the smaller sample.

My description of these two approaches, direct reading and collab- orative reading, is intended as not only a contrast with the computa- tional and probabilistic methods that will be introduced shortly; it is also a reminder that there are many ways of exploring a very large corpus. Researchers should not be intimidated by quantity. Even a million books could be studied by gathering a large random sample and using collabora- tive reading.

Machine Reading: Latent Dirichlet Allocation and Topic Models

Other ways of reading a very large collection of texts exist. A range of alternative approaches might be labeled, following N. K. Hayles, "machine reading."[9] In this section, I will introduce one of these alterna- tives, known informally as a *topic model*.

Like any abstraction, the vector space model obscures important aspects of texts, word order chief among them—for example, "the child ate the fish" and "the fish ate the child" are indistinguishable. It fails spectacularly when confronted with polysemy: "Mann" in "Ein junger Mann" is counted the same as the "Mann" in "Thomas Mann." And many measures used to compare word count vectors are maddeningly opaque. For example, while it is tempting to characterize cosine distance as a measure of similarity, this similarity has no interpretation familiar to human readers. And as a practical matter, in cases where one is dealing with roughly comparable texts, experiments have shown that cosine dis- tance and related measures are only loosely correlated with human judg- ments of similarity.[16]

Another objection to the vector space model is that readers often do not care about individual words *per se*; rather, they are interested in *groups of related words*. For example, if we really wanted to capture how much each chapter of *Effi Briest* featured Effi, we would want to consider all the words associated with her. She is called "Effi" by her parents and Innstetten, but she is called "gnädige Frau" by others. We would also be interested in the possessive form "Effis" along with the inflected forms of "gnädige Frau." These are all distinct vocabulary items in the vector space

model. Similarly, with our corpus of journal articles, if we were interested in identifying the proportion of articles devoted to a certain topic, such as the study of German folktales, we would be interested in a *set* of words, such as "tale," "tales," "fairy," "grimm," "folk," "wilhelm," and "broth- ers." If we were interested in the rise of feminist criticism, we would be concerned with tracking the occurrence of a cluster of words, such as "women," "woman," "male," "feminist," "gender," "patriarchy," and "social." Whether we are working with the chapters of a novel or with journal articles, it would be convenient to relax the vector space model somewhat and instead represent texts in terms of these distinctive constel- lations of words.

Remarkably, human readers need not specify which words belong to these clusters of words. Given a large corpus of texts, these groups of related words can often be *inferred* from their patterns of occurrence alone. In a limited sense, the data—here, the corpus—can "speak for itself." Making use of a *topic model* is one way of achieving this feat.

Latent Dirichlet Allocation and Topic Models

Topic model is an informal label for a member of a family of probabilistic models developed over the last ten years. These models trace their roots to a model described in 2003 by David Blei, Andrew Ng, and Michael Jordan.[17] The authors named this model Latent Dirichlet Allocation, or LDA. *Latent* refers to the model's assumption that the aforementioned clusters of words exist and are responsible in a specific sense for the word frequencies observed in the corpus. As these groups of words are themselves hidden, their distribution in the corpus needs to be inferred. *Dirichlet* refers to the probability distribution that does this work. The distribution is named after the nineteenth-century German mathemati- cian Peter Gustav Lejeune Dirichlet (1805–59).[18] The name *topic model* was retrospective. In practice, the model successfully finds groups of related words in a large corpus of texts—groups of words that readers felt comfortable calling *topics*.[19] Strictly speaking, these topics are prob- ability distributions over the unique words (vocabulary) of the corpus; those words to which the distributions assign the highest probability are those I will refer to as *associated* or *linked* with the topic. While new topic models have appeared in the intervening years, I will use LDA to model the journal article corpus.[20]

To understand how LDA works it is easiest to start with the end result.[21] LDA delivers a representation of each document in terms of topic shares or proportions. For example, assuming that thirty topics are latent in the corpus, the words in the article by Catherine Dollard, "The *alte Jungfer* as New Deviant: Representation, Sex, and the Single Woman in Imperial Germany," are associated with topics in the following

proportions: 47 percent topic 25, 17 percent topic 19, and 9 percent topic 20 (with 27 percent distributed with smaller shares over the remain- ing 27 topics; fig. 3.5). The plurality of the words is associated with topic 25, which in turn is characterized by its assigning high probability to observing the following words: "women," "female," "woman," "male," "sexual," "feminist," "social," "gender," "family," and "mother."

How does LDA arrive at this representation? Should readers trust its description of articles in the corpus? The first question has a ready answer. LDA and other topic models add an interpretive layer on top of the vec- tor space model. These models look at word frequencies through the lens of probability, permitting considerable flexibility in the interpreta- tion of the counts. I work through the details of a simple topic model in an online appendix to this chapter.[22] Recall that when we are thinking in terms of cosine distance (which is not probabilistic), observing that two documents share a word (e.g., "weimar") counts immediately as evi- dence of similarity. With probability added, judgment of similarity can be postponed and made in the context of other evidence (i.e., other shared words). This flexibility is advantageous when we are dealing with the fact of polysemy in human language—a single word frequently has a diversity of meanings. For example, consider two articles that both use "weimar," one concerning Goethe (who lived in this city) and one about the Weimar Republic. Seeing the word "weimar" in both documents should *not neces- sarily* count as evidence that the two documents concern similar subjects.

The addition of probability to the model permits the association of the word "weimar" with two different topics.

Should we trust that the description of documents in terms of top- ics corresponds at all with what our judgments would have been, had we read the 22,198 articles? The titles of journal articles provide a validation of the model. Recall that the topic model only uses the text of the article; words in the title are given no special status. Verifying that what the topic

	share
Topic 25	.47
Topic 19	.17
Topic 20	.09

Dollard, Catherine. "The *alte Jungfer* as New Deviant: Representation, Sex, and the Single Woman in Imperial Germany," *German Studies Review* 29 (Feb 2006): 107-26.

	top words
Topic 25	women female woman male sexual feminist social gender family
Topic 19	german political social history austrian national studies germany
Topic 20	life time people death love little story world father day left

Figure 3.5. Catherine Dollard's *German Studies Review* article viewed in terms of prominent topics. Shares and words are based on a topic model (LDA) with thirty topics. Considered separately, each of the remaining topics contributes less than 0.05.